



ASHOKA
UNIVERSITY



Centre for
Social and
Behaviour
Change

COVID-19 vaccine hesitancy in rural India: Survey evidence from Bihar and Uttar Pradesh

PRE-ANALYSIS PLAN

MARCH 2022

PRIMARY INVESTIGATORS

Dr Sharon Barnhardt

Dr Pavan Mamidi

Dr Sneha Shashidhara

CONTRIBUTORS

Selva Swetha

Shardul Vaidya

[This project is officially documented here, on the AEA RCT Registry.](#)

1 Introduction

1.1 Abstract

We intend to conduct a large in-person survey in rural India to understand the determinants of vaccine hesitancy amongst a primarily unvaccinated sample. Our study will be conducted in Uttar Pradesh and Bihar, with an exclusive focus on data collection from rural areas. Our survey will empirically estimate demand for the vaccine, the determinants of this demand or the lack thereof, and hurdles due to supply-side issues. We will also collect information about knowledge related to COVID-19 and COVID-19 vaccines, adherence to COVID-19 protective measures, media consumption and demographics to serve as covariates in our analysis. In addition to the survey, we also intend to test behaviourally informed communications interventions that we hypothesise could improve vaccine agency amongst our sample. Our sample size will be about 6000 respondents and roughly representative of the rural population of each state.

1.2 Introduction

There is a fair amount of anecdotal evidence that there is some hesitancy to avail a COVID-19 vaccine in India.¹²³⁴ Surprisingly, there is not enough quantitative evidence to establish the level of hesitancy and, more importantly, the beliefs and barriers behind the hesitancy, especially in the context of rural India. Large online surveys across different states in India have focused on getting rapid measurements of hesitancy levels.⁵⁶⁷ They use a non-exhaustive, shortlist of reasons/barriers relating to vaccine hesitancy for

¹ Singh, R. K. S. (2021, June 21). Vaccine hesitancy puts India's gains against virus at risk. AP News. <https://apnews.com/article/India-science-coronavirus-pandemic-health-8dd07a1f6bb56f4352130307f843458f>

² Mishra, A. (2021, June 23). Vaccine hesitancy, Covid myths rife in UP's Raebareli. India Today. <https://www.indiatoday.in/coronavirus-outbreak/story/vaccine-hesitancy-covid-myths-up-raebareli-ground-report-1818577-2021-06-23>

³ Sinha, B. C. (2021, July 1). Covid India: Women in rural Bihar hesitant to take vaccines. BBC News. <https://www.bbc.com/news/world-asia-india-57551345>

⁴ Mashal, M., & Kumar, H. (2021, July 11). One Village Quelled the Virus. The Next Was Overrun. It's a Bad Sign for India. The New York Times. <https://www.nytimes.com/2021/07/11/world/asia/covid-india.html>

⁵ IANS 2021

⁶ Prashnam 2021

⁷ LocalCircles 2021

respondents to choose from, and cannot report adequate demographic information about respondents.

Facebook has been deploying surveys in association with the University of Maryland and Carnegie Mellon University since December 2020, measuring a host of COVID-19 related outcomes, including symptoms, behaviour, vaccine attitudes and uptake. While the Facebook data collection exercise is relatively concise in investigating barriers and attitudes, its sampling frame consists exclusively of Facebook users. It is less than ideal for making policy recommendations for less privileged subgroups, given the mismatch between the demography of Facebook users and the Indian population. For instance, a 2019 UNDP report on Social Media Usage in India states that Facebook usage is primarily an urban-male-youth phenomenon, casting doubts over how well the results from the survey mirror attitudes of marginalised groups (e.g., women, low-income groups, senior citizens) in rural India, which have had historically lower rates of access to the internet and technology.⁸ This, more broadly, is the limitation of most vaccine hesitancy studies conducted via online surveys. They are not representative of the digitally marginalised population that lack access to smartphones and the internet. A study by PCI is the closest to a representative measurement we have for rural India (in particular, Bihar).⁹ They utilise a random sampling method, conducting in-person interviews to collect hesitancy attitudes and measure the effect of a list of barriers on perceived self-efficacy to get vaccinated. However, data was collected before the second wave of COVID hit India, and they have a fairly small sample size (N=872).

The urgent need is to design and field an in-person survey in rural India that targets populations least represented in the online surveys, i.e. female (48% nationally)¹⁰, rural home (65% nationally)¹¹, above 60 years old (8.6%)¹⁰ and illiterate (22% at a national level, 15% for males and 29% for females, according to NSSO).¹² We propose rural UP (160 villages, 20 districts) and rural Bihar (96 villages, 12 districts) to be the ideal locations for the study, given the large

⁸ UNITED NATIONS VOLUNTEERS INDIA OFFICE. 2019. "Social Media For Youth & Civic Engagement In India". United Nations Development Programme. <https://www.in.undp.org/content/dam/india/docs/UNV/Publications/SOCIAL%20MEDIA%20REPORT.pdf>.

⁹ Project Concern International. 2021. COVID-19 Vaccine Eagerness, Hesitancy and Refusal in Rural Bihar: A Barrier Analysis (BA) Study. Project Concern International.

¹⁰ Census of India. 2011. Office of the Registrar General & Census Commissioner, India. India

¹¹ World bank. World Bank staff estimates based on the United Nations Population Division's World Urbanization Prospects: 2018 Revision.

¹² National Sample Survey Report No. 585. 2017-18. Ministry of Statistics and Programme Implementation.

rural population (~248 million)¹³, low digital access (e.g. in rural Bihar, 39% men and only 17% women say they have ever used the internet¹⁴¹⁵).

This document outlines our proposed in-person survey using the 2011 Census as our sampling frame using a proportional sampling method. We will measure the barriers critical to vaccination, both demand and supply-side, knowledge and attitudes towards COVID-19 and its vaccination. We will also test 6 different interventions that we hypothesise will reduce vaccine hesitancy, delivered at the end of the survey through short videos. The effectiveness of these interventions will be measured through intention to get vaccinated, interest and trust placed in the videos.

¹³ Census of India. 2009. Office of the Registrar General & Census Commissioner, India. India

¹⁴ International Institute for Population Sciences. (2021). National Family Health Survey (NFHS-5), 2019-20: India. Mumbai, India: International Institute for Population Sciences.

¹⁵ Coronavirus in India: Latest Map and Case Count. (n.d.). COVID19INDIA.ORG. Retrieved July 12, 2021, from <https://www.covid19india.org>

2 Methods

2.1 Experimental Design Overview

We plan to conduct the survey using the proportional sampling method (based on the 2011 Census) in 20 districts of Uttar Pradesh and 12 districts of Bihar. The list of districts is shown in Table 1. Within these districts, eight villages will be randomly selected for data collection. Each of these villages is sampled entirely, ensuring proportionate sampling. Participants are first screened on vaccination status. All unvaccinated participants will proceed to the survey, and 10% of participants with one dose and 10% of participants with two doses will complete the survey.

All vaccinated participants who complete the survey will be assigned to 3 arms: Control, T1, and T2. All unvaccinated participants who complete the survey will be assigned to 7 arms: Control, T1', T2', T3, T4, T5, T6.

Table 1: List of Districts in Both States

Sr. No.	State	District
1	Uttar Pradesh	Sitapur
2	Uttar Pradesh	Sultanpur
3	Uttar Pradesh	Lakhimpur Kheri
4	Uttar Pradesh	Hardoi
5	Uttar Pradesh	Kushinagar
6	Uttar Pradesh	Gonda
7	Uttar Pradesh	Bahraich
8	Uttar Pradesh	Moradabad

9	Uttar Pradesh	Raebareli
10	Uttar Pradesh	Badaun
11	Uttar Pradesh	Muzaffarnagar
12	Uttar Pradesh	Barabanki
13	Uttar Pradesh	Deoria
14	Uttar Pradesh	Bijnour
15	Uttar Pradesh	Bulandshahr
16	Uttar Pradesh	Maharajganj
17	Uttar Pradesh	Aligarh
18	Uttar Pradesh	Saharanpur
19	Uttar Pradesh	Siddharthnagar
20	Uttar Pradesh	Fatehpur
21	Bihar	Muzaffarpur
22	Bihar	Madhubani
23	Bihar	Samastipur
24	Bihar	West Champaran

25	Bihar	Vaishali
26	Bihar	Sitamarhi
27	Bihar	Araria
28	Bihar	Begusarai
29	Bihar	Bhojpur
30	Bihar	Madhepura
31	Bihar	Jamui
32	Bihar	Khagaria

2.2 Sample Identification and Recruitment

Our sample will consist of rural populations in Uttar Pradesh and Bihar. We chose 20 districts in Uttar Pradesh and 12 districts in Bihar using both absolute rural population estimates from the 2011 census and the percentage of the population that have had at least one dose of the COVID-19 vaccine from the COWIN dashboard¹⁶. We selected half of the districts with the largest rural populations in both States (36 in Uttar Pradesh, 19 in Bihar). Among these, we selected the top 20 and 12 districts in Uttar Pradesh and Bihar, respectively, with the least population having had at least one dose of the COVID-19 vaccine. This ensures we maximise the possibility of identifying rural unvaccinated people in these districts. Within each district, we select eight villages and five more buffer villages to be visited if we do not reach our target numbers.

Enumerators will start at one end of the village and visit every alternate house in the village. They will request to speak to the male head of the household in

¹⁶ <https://dashboard.cowin.gov.in/>

one house and the female head of the household in the other house and continue alternatively to maintain gender balance in the sample. They will continue until we reach our target (~25 people) in the village.

2.3 Power Analysis

We intend to recruit around 6,000 subjects for the study, which allows us to calculate sample estimates for vaccine hesitancy outcomes for the ~247 million rural population of UP and Bihar with a 99% confidence interval and margin of error of about 1.7%.

2.4 Treatment Arms

At the end of the survey, participants will watch short videos that target COVID-19 vaccine hesitancy, hesitancy concerning getting their children vaccinated for COVID-19 and mask-wearing behaviour. All vaccinated participants (with either 1 or 2 doses) will be randomly assigned to one of three arms, one control arm and two treatment arms. Details of these arms are shown in Table 2

Table 2: Description of the experiment arms for the vaccinated participants

Sr. no.	Experiment Arm	Description	Time
1	Control	A UN sponsored animated video describing the role of forests in climate change.	3 min 7 sec
2	T1 - Emotional Approach: Child vaccination	Is a conversation between a brother and a sister where the brother, using a story of a family where children get COVID-19, convinces his sister to vaccinate her children. The tone is emotional and touches upon severe disease in children and cost of hospitalisation, and the child vaccine being half the dose.	3 min 16 sec

3	T2 - Rational Approach: Child vaccination	Is a conversation between a brother and a sister where the brother convinces his sister to vaccinate her children using various facts and his expertise and experience as a front-line health worker. The tone is rational and touches upon severe disease in children and cost of hospitalisation, and the child vaccine being half the dose.	2 min 25 sec
---	---	--	--------------

All the unvaccinated participants will be randomly assigned to 7 experimental arms, one control arm and six experimental arms. Details of these are shown in Table 3. T1 and T2 are the same as shown to the vaccinated participants with one change. The video explicitly mentions that the adults are vaccinated and discusses children’s vaccination for the vaccinated participants. In contrast, that line in each of the videos was removed for the unvaccinated participants. In this document, these videos will be denoted as T1' and T2'.

Table 3: Description of the experiment arms for the unvaccinated participants

Sr. no.	Experiment Arm	Description	Time
1	Control	A UN sponsored animated video describing the role of forests in climate change.	3 min 7 sec
2	T1' - Emotional Approach: Child vaccination	Is a conversation between a brother and a sister where the brother, using a story of a family where children get COVID-19, convinces his sister to vaccinate her children. The tone is emotional and touches upon severe disease in children and the cost of hospitalisation—the child vaccine	3 min 16 sec

		being half the dose—no explicit mention of the adults being vaccinated.	
3	T2' - Rational Approach: Child vaccination	Is a conversation between a brother and a sister where the brother convinces his sister to vaccinate her children using various facts and his expertise and experience as a front-line health worker. The tone is rational and touches upon severe disease in children and the cost of hospitalisation—the child vaccine being half the dose—no explicit mention of the adults being vaccinated.	2 min 25 sec
4	T3 - Fear induce approach: Adult Vaccination	A short video inducing fear about COVID-19 by providing statistics of infection and death. It also mentions hospitalisation and intubation, along with relevant visuals.	0 min 48 sec
5	T4 - Rational Approach: Adult Vaccination	Is a short video that compares COVID-19 to other fatal diseases and discusses how unlike them, severe COVID-19 is easily preventable by getting vaccinated; along with relevant visuals	0 min 57 sec
6	T5 - Emotional Approach: Adult vaccination	Is an animated video that shows the journeys of two women, one vaccinated and another not. The vaccinated woman's family goes on to prosper, while the other woman dies due to infection destroying future prospects of the family and children.	2 min 42 sec

7	T6 - Emotional approach” Mask-wearing	Is an animated video that shows the journey of a family where one unmasked meeting leads to an older woman dying of COVID-19	2 min 29 sec
---	---------------------------------------	--	--------------

2.5 Randomisation

The recruitment of the participants for the survey does not include randomisation. All unvaccinated participants and 10% of participants with one dose and two vaccination doses pass the screening and complete the entire survey. For the Intervention videos, the participants are randomly assigned to one of 3 arms (vaccinated participants) or one of 7 arms (unvaccinated participants).

2.6 Data Collection

Enumerators from the NYAS agency were hired to administer the in-person survey on licensed software Survey CTO on their offline app. All adults in the villages will be invited to participate in the study and asked screening questions after obtaining consent. Enumerators conduct the survey in-person and type the answers into the app on smartphones simultaneously. The survey and interventions are in Hindi. 10% of all interviews will be recorded for quality checks. The duration of the survey is around 30 mins. The treatment videos and the following question take another 10 mins. Participants will not be compensated for their time. Only complete surveys will be used for analysis, and no participants with partial surveys will be recontacted to resume the survey. While the enumerators are familiar with the broad outline of the study, they are not made aware of the details of the treatment groups to reduce potential bias.

2.7 Experimental Flow

Step 1: Enumerators recruit participants in each village according to the protocol described above (2.2 Sample Identification and Recruitment). Suppose the vaccinated participants' target in that village has been met, the enumerators will informally ask participants their vaccination status and only proceed to the screening form if they are unvaccinated. This allows the vaccinated participants to be spaced out across all villages and districts.

Step 2: Enumerators will administer the informed consent and screening questions. Based on their vaccination status, the participants are screened in or screened out.

Step 3: Enumerators hired through an external agency conduct in-person surveys. Answers are marked on SurveyCTO offline app on smartphones.

Step 4: Deployment of treatment - Vaccinated participants are randomly assigned to one of 3 groups, and unvaccinated participants are assigned to one of 7 groups. All participants will watch a short video regarding COVID-19 vaccination or mask-wearing and answer a few questions after.

1. The screening section has questions on vaccination status and basic demographics. If respondents get screened-in, they are asked questions about their vaccination intention (for unvaccinated respondents), anticipated and experienced barriers to vaccination, knowledge about the vaccine, protective behaviour, trust in information sources, attitudes towards vulnerable populations about the COVID-19 vaccine, risk perceptions and demographics.
2. Our survey has two questions where we intend to collect open-ended audio responses. For these questions, the surveyor will read out the question and request the respondent to speak into the smartphone to record their response.

2.8 Pilot data

The entire survey and interventions will be tested, and each of the enumerators doing the data collection will collect two complete surveys. This is to ensure the data quality and address any concerns in the field.

2.9 Backcheck

To ensure the quality of the data, we will conduct backcheck interviews. 10% of the participants will be chosen for an additional short survey across all experiment arms and enumerators. A week after their primary survey, these surveys will be conducted by a separate set of enumerators via a phone call. It will include a few questions about the previous survey length and comfort rating, their vaccination intention, a few questions about COVID-19 related knowledge, and demographics questions.

3 Outcome Variables

We primarily want to investigate the barriers to vaccination faced by rural populations. We will measure the main barriers to vaccination in a free-form question and specifically about ten barriers to vaccination. We will also measure their intention to get vaccinated and when they plan to get vaccinated. In addition, we will measure knowledge about vaccination, concern about side-effects, risk perception of COVID-19, COVID appropriate practices. We will also measure critical demographics, media consumption trends, and trust in news sources. Details of the outcome variables measured in the survey are shown in Table 4.

At the end of the survey, participants will watch short intervention videos, and we will measure their intention to get vaccinated and their willingness to share this information. In addition, we will measure how engaging and how trustworthy the participants found these videos. Details of the outcome variables measured in the survey are shown in Table 5.

Table 4: Description of Outcome Variables in the survey

Outcome Variable	Description¹⁷	Outcome Measures
Intention to get vaccinated	A self-reported measure of whether they definitely, probably, probably not, and definitely not get vaccinated	Variable Type: Ordinal (0, 1, 2, 3)
Timeline for vaccination	A self-reported measure of whether they want to get vaccinated immediately, in the next 30 days, 1-3 months, 3-6 months, 6-12 months, after 12 months, or never.	Variable Type: Ordinal (0, 1, 2, 3, 4, 5, 6)

¹⁷ Please refer to the [survey](#) instrument for further details on all outcome measures across tables.

	(regarding the first dose for the unvaccinated; and second dose for those who have had one dose)	
Main Barriers for Vaccination	A self-reported reason for not getting vaccinated is asked as a free-form question coded into one of 36 reasons. Another question probes whether each of ten specific reasons for vaccine hesitancy applies to them (or applied to them before they got vaccinated).	Variable Type: Categorical (0, 1, 2, 3 to 36)
Problems faced during vaccination	A self-reported measure of problems faced during vaccination	Variable Type: Categorical (0, 1, 2, 3 to 16)
Vaccine Knowledge	Three questions on the function of the vaccine.	The number of correct responses. Variable type: Numerical (0-3)
Attitude towards vaccine compulsion	Self-reported measures using Likert scale rating (1-5) from complete agreement to a disagreement of vaccine compulsion and making vaccines compulsory for children to attend school.	Number with a Likert scale of '3' and above. Variable: Numerical (0, 1, 2)

Value Perception of Second Dose	Self-reported measures using Likert scale rating (1-5) from completely important to completely unimportant	Variable Type: Ordinal (0, 1, 2, 3,4)
Intentionally avoid contact	Self-reported measures using Likert scale rating (1-4) from all of the time so none of the time	Variable Type: Ordinal (0, 1, 2, 3)
Perception of the effectiveness of COVID appropriate behaviour	Self-reported measures using Likert scale rating (1-4) from very effective to not effective at all about social distancing and mask-wearing	Number with a Likert scale of '2' and above. Variable: Numerical (0, 1, 2)
Misinformation	Self-reported measures using Likert scale rating (1-5) from definitely true to false on not wearing masks after vaccination and children not getting infected.	Number with a Likert scale of '3' and above. Variable: Numerical (0, 1, 2)
Trust in channels/Sources	Self-reported measures of which channels and sources they trust wrt to COVID-19 news	The total number of channels/sources trusted. Variable type: Numerical (0-13)
Likelihood of vaccinating children	Self-reported measures using Likert scale rating (1-5) from completely likely to not likely at all	Variable Type: Ordinal (0, 1, 2, 3, 4)

<p>Risk perception of vaccine for people with other vulnerabilities</p>	<p>Self-reported measures about recommending the vaccine if the person is either pregnant or breastfeeding, diabetic, has cancer of TB</p>	<p>The sum of 'Yes' responses' Variable type: Numerical (0-5)</p>
<p>Infection risk perception</p>	<p>Self-reported measures using Likert scale rating (1-4) on worry about catching COVID and worry about getting the severe version during the second wave</p>	<p>Number with a Likert scale of '3' and above. Variable: Numerical (0, 1, 2)</p>
<p>Demographics</p>	<p>Age, gender, religion, caste, employment, support of current Central Govt., monthly household income per member, highest education level, number of adults in the house that are vaccinated, use of WhatsApp, toilet in the house, access to fresh drinking water, ever tested for COVID-19, time to the nearest medical centre, previous chronic health condition, and chew paan (betelnuts and leaves), smoke or drink liquor.</p>	<p>Age - Continuous variable Gender - Binary variable Religion and caste combined - Categorical variable (1-9) Employment - Categorical variable (1-5) Income - Log of household income per person. Continuous variable Education - Ordinal variable (1-9) No. of vaccinated adults in the house - Continuous variable Use of WhatsApp - Binary variable Support of Central Govt. - Binary variable Access to sanitation -</p>

		<p>number of 'yes' for access to fresh drinking water and toilet - Numerical (0, 1,2)</p> <p>Time to medical centre - Ordinal variable (0,1,2,3)</p> <p>Health conditions - Number of 'yes' to 7 common conditions - Numerical (0, 1,2,3,4,5,6,7)</p> <p>Paan, tobacco or liquor - Binary variable</p>
Vaccination status of children	<p>How many doses of COVID-19 vaccine have the children been given</p> <p>Children on schedule for routine vaccinations</p> <p>[for select participants with children]</p>	<p>0, 1, some 1 and some 2, 2 dose</p> <p>Variable Type: Ordinal (0, 1, 2, 3)</p> <p>Variable: Binary</p>

Table 5: The outcome variables after the intervention/control videos:

Outcome Variable	Description	Outcome Measures
Intention to get vaccinated	A self-reported measure of whether they definitely, probably, probably not, and definitely not get vaccinated	<p>Variable Type: Ordinal (0, 1, 2, 3)</p> <p>A difference variable is created by subtracting the last option from the current option of the same question.</p>

		Variable type: Numerical (-3 to +3)
Willingness to share the video	Revealed preference measure asking if the participants are willing to share a contact number of a friend/family member to whom the intervention video will be shared as coming from them	Variable type: Binary
Found the video engaging	Self-reported measures using Likert scale rating (1-5) from thoroughly engaging to not engaging at all	Variable Type: Ordinal (0, 1, 2, 3,4)
Found the video trustworthy	Self-reported measures using Likert scale rating (1-5) from completely trustworthy to not trustworthy at all	Variable Type: Ordinal (0, 1, 2, 3,4)
Manipulation check question	One question per treatment arm specific to that video	Variable type: Binary
For Control, T1, T2, T1', T2': Likelihood of getting children vaccinated	Self-reported measures using Likert scale rating (1-5) from Very likely to not likely at all	Variable Type: Ordinal (0, 1, 2, 3, 4) A difference variable is created by subtracting the last option from the current option of the same question.

		Variable type: Numerical (-4 to +4)
For Control, T6: Likelihood of wearing masks	Self-reported measures using Likert scale rating (1-5) from Very likely to not likely at all	Variable Type: Ordinal (0, 1, 2, 3, 4)

4. Analysis

4.1 Model Specifications:

Ordered Logistic Regression will be used for ordinal outcomes variables and multinomial logit regression for the categorical outcome variables.

Ordinary Least Squares regression will be used for discrete and continuous numerical outcome variables.

For each outcome measure, we will estimate two models, one controlling for demographic covariates and one without.

Fourteen outcome measures (Table 4) for a binary vaccination status variable (unvaccinated: 80% of the sample, vaccinated: 20% of the sample) times two models make 28 hypothesis tests. Thus, we will use multiple hypothesis testing adjustments with pFDR and the q-value¹⁸.

We will estimate models for the form:

M1: $Y \sim \text{vaccination_status} + \text{demographic_covariates} + \text{error}$

M2: $Y \sim \text{vaccination_status} + \text{error}$

where Y is any outcome measure from Table 4

vaccination_status = dummy variable, 1 for vaccinated (1 or 2 doses) and 0 for unvaccinated.

¹⁸ Storey, JD. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64 (3): 479–498. <https://doi.org/10.1111%2F1467-9868.00346>

Six outcome measures (Table 5) for two treatments (Only vaccinated participants: 20% of the sample) times two models make 24 hypothesis tests. Thus, we will use multiple hypothesis testing adjustments with pFDR and the q-value¹⁸.

M3: $Y \sim \text{treatment_assignment} + \text{demographic_covariates} + \text{error}$

M4: $Y \sim \text{treatment_assignment} + \text{error}$

where Y is any outcome measure from Table 5

treatment_assignment = dummy variable, 1 for treatment and 0 for control.

We have five outcome measures (Table 5), six treatments (Only unvaccinated participants: 80% of the sample) times two models. For T1, T2, and T6, we have an additional outcome variable, making a total of 66 hypothesis tests. Thus, we will use multiple hypothesis testing adjustments with pFDR and the q-value¹⁸.

M5: $Y \sim \text{treatment_assignment} + \text{demographic_covariates} + \text{error}$

M6: $Y \sim \text{treatment_assignment} + \text{error}$

where Y is any outcome measure from Table 5

treatment_assignment = dummy variable, 1 for treatment and 0 for control.

All analysis, including randomisation and data checks, will be conducted using custom-made MATLAB (The MathWorks, Inc) scripts in R (R Core Team, 2014)¹⁹.

4.2 Randomisation Check

Treatment status is the only difference between the treatment and control groups in a randomised control study. On average, all other characteristics of treatment and control group members, including demographics, should be balanced. Treatment effect estimates could be biased if there is an imbalance across the groups despite the randomisation process. We will check for balance between treatment and control groups for key covariates.

¹⁹ R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>

M7: $X \sim \text{treatment_assignment} + \text{error}$

X are the different demographic covariate.